Suhwan Choi

◉ Seoul, South Korea ✉ milkclouds00@gmail.com in Suhwan Choi 🔗 milkclouds.work ⌂ milkclouds

Summary

I am an undergraduate student majoring in Physics and Computer Science at Seoul National University. My main area of interest is approximating and imitating human behavior and intelligence in multmodal modalities, utilizing end-to-end architectures and scalable training suites.

Experience

Principal Researcher *Feb 2024 – Present*
Maum.ai

- Founded autonomous robotics research division as the first researcher, leading strategic decisions and team expansion to 10 researchers.
- Contributed as first author to majority of research projects in robotic navigation and embodied ai.
- Led CORE: Slurm-based DGX Cluster construction project (96 H100 GPUs, 12 nodes). [Blog] ↗
- Implemented company-wide Notion workspace enhancing productivity, streamlining workflows. [Template] ↗

Machine Learning Engineer Intern *July 2023 – Jan 2024*
Hyperconnect

- Worked on diffusion-based personalized profile image generation for real-world applications. The output images were conditioned not only on prompts but also on the face in the input image.

Publications

D2E: Scaling Vision-Action Pretraining on Desktop Data for Transfer to Embodied AI [Website] ↗ — 1st author

Suhwan Choi*, Jaeyoon Jung*, Haebin Seong*, Minchan Kim, Minyeong Kim, Yongjun Cho, Yoonshik Kim, Yubeen Park, Youngjae Yu†, Yunsung Lee†

Under Review

CANVAS: Commonsense-Aware Navigation System for Intuitive Human-Robot Interaction [Website] ↗ — 1st author

Suhwan Choi*, Yongjun Cho*, Minchan Kim*, Jaeyoon Jung*, Myunchul Joe, Yubeen Park, Minseo Kim, Sungwoong Kim, Sungjae Lee, Hwiseong Park, Jiwan Chung, Youngjae Yu†

ICRA 2025
Workshop on Open-World Agents at NeurIPS 2024 **(Outstanding paper Awards, 3%)**

Revisiting Residual Connections: Orthogonal Updates for Stable and Efficient Deep Networks — 4th author

Giyeong Oh, Woohyun Cho, Siyeol Kim, **Suhwan Choi**, Youngjae Yu†

NeurIPS 2025

ESREAL: Exploiting Semantic Reconstruction to Mitigate Hallucinations in Vision-Language Models — 2nd author

Minchan Kim*, Minyeong Kim*, Junik Bae*, **Suhwan Choi**, Sungkyung Kim, Buru Chang†

ECCV 2024

Education

Seoul National University *2021 – Present*
B.S in Physics & Computer Science and Engineering

Awards

QHack Coding Challenge [link] ↗   *2023 and 2024*
- Ranked **4th/793** teams in 2023, Ranked **3rd/618** teams in 2024.
- This is a contest implementing quantum algorithms, quantum machine learning, quantum chemistry, and brain-teasing puzzles.

2023 Quantum Hackathon [link] ↗   *2023*
- **1st place, Minister of Science and ICT Award**
- Topic: Utilizing symmetry to solve variational quantum algorithm (quantum machine learning) efficiently.

NAVER CLOVA AI RUSH 2022 [link] ↗   *July 2022 – Sept 2022*
- **3rd place** on Landmark Detection (3,000,000 KRW)
- **2nd place** on Shopping User Embedding Extraction, Classification (7,000,000 KRW)

Google Codejam 2022   *2022*
- Round 3, 546th (awarded T-Shirt).

Activities

Open Source Contributions
- **Open World Agents** ↗**:** Core contributor and maintainer with 150+ merged PRs. Built comprehensive multimodal desktop agent framework including optimized data collection tool, standardized efficient data format, multimedia data processing pipelines, dataset management, agent training and also Python packaging, CI/CD infrastructure.
- **LinkedIn Liger-Kernel** ↗ **(5.8k stars):** Implemented InternVL3 support for memory-efficient kernel optimization. Added multimodal model support enabling efficient training with reduced memory footprint.
- **HuggingFace (235k+ total stars):** 5 merged PRs and 15+ issues across core ML libraries. Enhanced ImageProcessor API consistency for multimodal models, improved PeftModel integration with proper warning systems, enabled batch generation in community SDXL pipelines, and identified critical bugs in multi-device training and tokenizer handling.
- **Open-mmlab:** Fixed critical deadlock issues in mmdetection (31.8k stars) and mmsegmentation (9.3k stars) for distributed multi-GPU environments. Implemented label visualization in mmdetection3d (6.1k stars).

Technical Blogging
- Writing tech blog milkclouds.work ↗ covering AI/ML research, robotics, industry analysis, and practical engineering insights

Deepest: SNU Deep Learning Society [Website] ↗   *Sept 2022 – Present*
- Participated in research projects: neural ODE study, RL-based game solving ↗, parallel computing, causal inference, voice generation

Software Maestro, Trainee   *2021*
- Studied 3D vision and implemented AR game for iPad using AR Foundation of Unity
- Zombie is evading sofa because iPad realized furnitures, and occlusion is also working. [link] ↗

Skills

Programming Languages: Highly proficient in Python / Experience with: C++, JAVA, HTML/CSS/JS, Golang, Rust, ...

ML/AI: PyTorch, HuggingFace (Transformers, Diffusers, LeRobot, ...), open-mmlab libraries, ROS

Infrastructure & DevOps: Linux, Slurm, Docker, Ansible, Ceph, GitHub Actions CI/CD, HPC infrastructure

Languages: Korean (Native), English